Linear Algebra

Samira Hossein Ghorban
s.hosseinghorban@ipm.ir

Fall, 2021

System Equations

We want to fit a plane $y = C + Dt + Ez$ to the four points

$$y = 3 \quad \text{at } (t, z) = (1, 1), \quad y = 6 \quad \text{at } (t, z) = (0, 3),$$

$$y = 5 \quad \text{at } (t, z) = (2, 1), \quad y = 0 \quad \text{at } (t, z) = (0, 0),$$

- $y = C + Dt + Ez$

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} C \\ D \\ E \end{bmatrix} = \begin{bmatrix} 3 \\ 5 \\ 6 \\ 0 \end{bmatrix}$$

- $C = 0,\ D =?,\ E = 2.$

Least Square

- Find C and D such that the following data are fitted by $b = C + tD$.

t	b	t	b
1.0	−1.945	3.2	0.764
1.2	−1.253	3.4	0.532
1.4	−1.140	3.6	1.073
1.6	−1.087	3.8	1.286
1.8	−0.760	4.0	1.502
2.0	−0.682	4.2	1.582
2.2	−0.424	4.4	1.993
2.4	−0.012	4.6	2.473
2.6	−0.190	4.8	2.503
2.8	0.452	5.0	2.322
3.0	0.337		

Least squares problems with several variables

- So, instead of one column and one unknown x, the matrix now has n columns. The number m of observations is still larger than the number n of unknowns.
- So it must be expected that $Ax = b$ will be inconsistent. Probably, there will not exist a choice of x that perfectly fits the data b.
- In other words, the vector b probably will not be a combination of the columns of A; it will be outside the column space.
- Solution: finding $\hat{x}$ which minimizes $E = \|b - Ax\|$.
- Equivalently, finding $\hat{x}$ such that the error vector $e = b - A\hat{x}$ be perpendicular to that space $C(A)$.

Least squares problems with several variables

- Finding $\hat{x}$ such that the error vector $e = b - A\hat{x}$ be perpendicular to that space $C(A)$ is equivalent to

$$A^T(b - A\hat{x}) = 0$$
$$A^T A\hat{x} = A^T b$$

- If $A^T A$ is invertible, then the best square estimation :

$$\hat{x} = (A^T A)^{-1} A^T b.$$

Weighted Least Squares

- Now suppose that the two observations are not trusted to the same degree.

- A simple least-squares problem is to estimate two observations $x = b_1$ and $x = b_2$.
- if $b_1 \neq b_2$, then we are faced with an inconsistent system of two equations in one unknown:

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} \begin{bmatrix} x \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}.$$

- Thus

$$E^2 = (x - b_1)^2 + (x - b_2)^2$$

- The best square estimation:

$$\hat{x} = \frac{b_1 + b_2}{2}.$$

Weighted Least Squares

- Now, suppose that the the value $x = b_1$ may be obtained from a more accurate source from a larger $x = b_2$.

- The simplest compromise is to attach different weights w_1 and w_2 and choose the $\hat{x}_W$ that minimizes the weighted sum of squares:

$$E^2 = w_1^2(x - b_1)^2 + w_2^2(x - b_2)^2.$$

- The best square estimation :

$$\hat{x} = \frac{w_1^2 b_1 + w_2^2 b_2}{w_1^2 + w_2^2}.$$

- Generally:

$$(A^T W^T W A)\hat{x}_W = A^T W^T W b.$$

Approximation

A best approximation

A **best approximation** to b by vectors in W is a vector $ainW$ such that

$$\|b - \hat{a}\| \leqslant \|b - a\|$$

for each vector $a \in W$. The vector $\hat{a} \in W$ is called a best approximation to b in W.

Approximation

Theorem

Let W be a subspace of an inner product space V and let b be a vector in V.

(i) *The vector $\hat{a}$ in W is a best approximation to b by vectors in W if and only if $b - \hat{a}$ is orthogonal to every vector in W.*

(ii) *If a best approximation to b by vectors in W exists, it is unique.*

(iii) *If W is finite-dimensional and $\{v_1, \ldots, v_n\}$ is any orthonormal basis for W, then the vector*

$$\hat{a} = \sum_{k=1}^{n} \langle b, v_k \rangle v_k$$

is the (unique) best approximation to b by vectors in W.

Inner product on function spaces

- Let V be a space of all continuous functions on the interval $[0, 2\pi]$.

- How can we measure their "length" in the space?
- The same idea of replacing summation by integration produces the inner product of two functions.
- Let V be a space of all continuous functions on the interval $[0, 2\pi]$ and $f, g \in V$, and define

$$\langle f, g \rangle = \int_0^{2\pi} f(x) g(x) \, dx.$$

- thus,

$$\| f \|^2 = \int_0^{2\pi} f(x)^2 \, dx.$$

Inner product on function spaces

- Consider $\sin(x), \cos(x) \in V$.

$$\|\sin(x)\|^2 = \int_0^{2\pi} \sin^2(x)\,dx = \frac{1}{2}\left(2\pi - \int_0^{2\pi} \cos 2x\right) = \pi.$$

$$\|\cos(x)\|^2 = \int_0^{2\pi} \cos^2(x)\,dx = \frac{1}{2}\left(2\pi + \int_0^{2\pi} \cos 2x\right) = \pi.$$

$$\langle \sin(x), \cos(x) \rangle = \int_0^{2\pi} \sin x \cos x dx = \frac{1}{2}\int_0^{2\pi} \sin 2x dx = 0$$

- Thus, the functions $\sin(x)$ and $\cos(x)$ are orthogonal. Thus, $\frac{1}{\sqrt{\pi}} \sin(x)$ and $\frac{1}{\sqrt{\pi}} \cos(x)$ are orthonormal.
- so, $\frac{m}{\sqrt{\pi}} \sin mx$ and $\frac{n}{\sqrt{\pi}} \cos nx$ are orthonormal.

Gram-Schmidt for Functions

- Are $\{1, x, x^2\}$ are orthogonal?
- we have

$$\langle 1, x\rangle = \int_{-1}^{1} x dx = 0.$$

- Let $q_1 = \frac{1}{\sqrt{\langle 1,1\rangle}}$ and $q_2 = \frac{x}{\sqrt{\langle x,x\rangle}}$.

$$Q_3 = x^2 - \frac{\left\langle x^2, 1\right\rangle}{\langle 1, 1\rangle}1 - \frac{\left\langle x^2, x\right\rangle}{\langle x, x\rangle}x = x^2 - \frac{\int_{-1}^{1} x^2\, dx}{\int_{-1}^{1} x dx} - 0 = x^2 - \frac{1}{3}$$

- The polynomials $1, x, x^2 - \frac{1}{3}$ are orthogonal to each other over the interval $-1 \leqslant x \leqslant 1$ they are called ***Legendre polynomials***.

Best Straight Line

- Suppose we want to approximate $y = x^5$ by a straight line $C + Dx$ between $x = 0$ and $x = 1$.
- There are at least three ways of finding that line.

1. Solve $\begin{bmatrix} 1 & x \end{bmatrix} \begin{bmatrix} C \\ D \end{bmatrix} = x^5$ by least squares.
2. Minimize
$$E^2 = \int_0^1 (x^5 - C - Dx)^2 \, dx.$$
3. Apply Gram-Schmidt!

Thank You!